REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 1992	Technical Memorandum

4. TITLE AND SUBTITLE
Development of a Machine Vision System for Automated Structural Assembly

5. FUNDING NUMBERS
WU 506-43-41-02

6. AUTHOR(S)
P. Daniel Sydow and Eric G. Cooper

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NASA Langley Research Center
Hampton, VA 23665-5225

8. PERFORMING ORGANIZATION REPORT NUMBER
L-16995

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
NASA TM-4366

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified–Unlimited

Subject Category 63

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Research is being conducted at the NASA Langley Research Center to develop a telerobotic assembly system designed to construct large space truss structures. This research program has been initiated within the past several years, and a ground-based test-bed has been developed to evaluate and expand the state of the art. Test-bed operations currently use predetermined ("taught") points for truss structural assembly. Total dependence on the use of taught points for joint receptacle capture and strut installation is neither robust nor reliable enough for space operations. Therefore, a machine vision sensor guidance system is being developed to locate and guide the robot to a passive target mounted on the truss joint receptacle. The vision system hardware includes a miniature video camera, passive targets mounted on the joint receptacles, target illumination hardware, and an image processing system. Discrimination of the target from background clutter is accomplished through standard digital processing techniques. Once the target is identified, a pose estimation algorithm is invoked to determine the location, in three-dimensional space, of the target relative to the robots end-effector. Preliminary test results of the vision system in the Automated Structural Assembly Laboratory with a range of lighting and background conditions indicate that it is fully capable of successfully identifying joint receptacle targets throughout the required operational range. Controlled optical bench test results indicate that the system can also provide the pose estimation accuracy required to define the target position.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Space trusses; Machine vision; Robotics; Automated assembly; Space structures	28
	16. PRICE CODE A03

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified		

NSN 7540-01-280-5500

Standard Form 298(Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

NASA-Langley, 1992

Introduction

Future space missions to and from planet Earth will probably require large truss structures to provide a stiff and stable platform for experimental measurements and observation antennas and telescopes. One method of obtaining such large structures is to construct them using an automated telerobotic system. Such a system could utilize either an astronaut or an Earth-based operator as an executive monitor who intervenes only when the automated system encounters a problem or requires assistance. This mode of operation, known as supervised autonomy, promises to accomplish large or complex assembly and construction tasks if the crew resources that are available on orbit are limited. Another advantage of supervised autonomy is that it can be monitored from the ground because it involves no time-critical, in-the-loop control functions as does a nonautomated telerobotic system.

Research currently is being conducted at the NASA Langley Research Center to develop a telerobotic automated structural assembly system designed to construct large space truss structures (ref. 1). This research program has been initiated within the past several years, and a ground-based test-bed has been developed to evaluate and expand the state of the art. Test-bed operations currently use predetermined ("taught") points for truss structural assembly. Total dependence on the use of taught points for joint receptacle capture and strut installation is neither robust nor reliable enough for space operations. Therefore, a machine vision sensor guidance system is being developed to locate and guide the robot to a passive target that is mounted on the truss joint receptacle. The intent of this proposed system is to provide sensor-based guidance that will permit both fully automated and operator-assisted structural assembly evaluation tests to be performed during the operation of the Automated Structural Assembly Laboratory. This paper describes the development and preliminary tests performed on a sensor guidance system that is based on machine vision techniques. System hardware and software development is presented, and verification tests in the assembly laboratory as well as controlled optical bench tests are presented and discussed. The authors acknowledge the support given by Brian M. Romansky for the collection of experimental data using the optical bench.

Assembly Operations and Vision System Requirements

Laboratory and Truss Hardware Description

A schematic of the Automated Structural Assembly Laboratory (ASAL) is depicted in figure 1(a), and a photograph of the actual hardware is shown in figure 1(b). This facility is a ground-based research tool to develop and evaluate assembly hardware concepts, construction techniques, software, and operator interface systems that are anticipated to be required for on-orbit assembly operations. The robot arm is a commercially available, electrically controlled, six degree-of-freedom industrial model that has been selected for the laboratory test operations because of its payload capacity, reach envelope, and positioning repeatability. No modifications to the robot other than those that are commercially available from the manufacturer have been made. The robot is mounted on an x-y Cartesian motion base that provides the translations to position the base of the robot anywhere in the support track area. This positioning can obtain an accuracy of 0.002 in. The truss is mounted to, and assembled on, a rotating motion base at the end of the translational base. In figure 1(a), θ is the angle at which the motion base is positioned. Both motion bases are designed to minimize positional errors that may be induced by the static deformations from the mass of the robot, an unbalanced asymmetric truss assembly, and the forces exerted by the robot during assembly.

The truss selected for the initial assembly test operations is a regular tetrahedral truss composed of 102 strut members (each approximately 6.56 ft (2 m) long). The size and configuration were chosen because they represent the support structures that are anticipated to be required for a number of planned or proposed missions. The truss strut members are connected by specially designed joints located near the nodes (ref. 2). Each node must be capable of connecting nine members (six that are in the plane of the top or bottom face of the truss and three that connect the top and bottom faces of the truss). The joints are located as close to the nodes as physically practical to accommodate structural requirements and to minimize the packaging volume of the truss strut members for launch. Locating the joints close to the nodes complicates the assembly operations because the robot end-effector must be small and capable of installing a member in a very confined region near the node.

Strut Installation

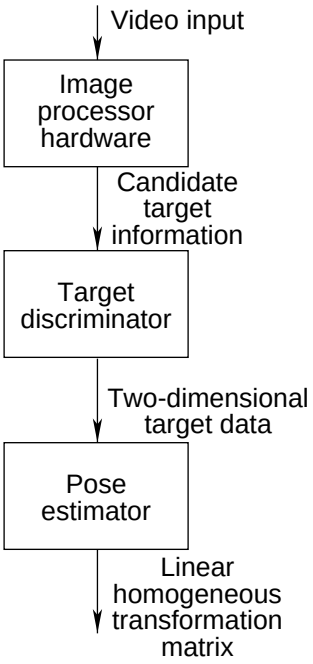
The current robot end-effector used in the truss assembly is a special-purpose device designed to be compatible with the truss joint. To insert the strut into the truss, the end-effector is moved to a taught position. At this location, the end-effector fingers are configured to align with the connector receptacle (if it is at any location within a cylindrical envelope that is 2 in. in diameter by 0.6 in. in length). Both ends of the strut are pushed forward by mechanisms (called platforms) which insert the connector into the joint receptacle. The strut is held in place while small gear-head motors lock the joints. The strut holders are then unlatched, the platforms retracted, and the end-effector fingers opened to complete the installation. Strut removal follows this procedure in reverse order. The end-effector is designed to permit operation either with a node preattached to either end of the strut or to insert a strut into nodes already in the structure.

The current telerobotic delivery of a strut from the canister for installation in the structure is made up of two phases. In the first phase, a path composed of several taught points guides the end-effector from a rest position above the canister to the insertion position where the fingers on the end-effector can grasp the joint receptacle. After the fingers have closed on the receptacles, the second phase begins. In this phase, a force and torque feedback system precisely aligns the end-effector with the aid of passive guidance features to eliminate all forces and torques that result from positioning errors. After positioning errors are reduced to a nominal level, the strut is totally installed by the end-effector. The telerobotic delivery of a strut from the canister for installation using a sensor-based guidance system is projected to consist of three phases. First, an automated path planner utilizes geometric information of the structure, robot arm, and carriage positions to guide the arm to a position. Second, the sensor system mounted on the end-effector is used to locate targets on the joint receptacle and to guide the arm into a position where the end-effector fingers can grasp the joint receptacle. Third, the force and torque feedback system precisely aligns the end-effector for strut insertion. Although a path-planning algorithm for the automated structural assembly system has not been evaluated, such a software tool is anticipated to be capable of guiding the arm to a point equal to the current operational requirements of the end-effector at a distance about 18 in. from the joint receptacle. This point is denoted as the vision approach point (VAP). The targets need to be specially designed and accurately fabricated and positioned on each joint receptacle to provide the information necessary for guiding the robot end-effector to the proper alignment position with the joint receptacle at 4 in. to 6 in. from the node for strut insertion and removal.

Sensors

Two alternative sensor technologies for locating the joint receptacle and guiding the arm were considered; these technologies include triangulation with infrared sensors and machine vision. The use of triangulation with infrared sensors in ASAL was rejected because initial tests revealed several potential drawbacks. First, the small field of view of a single infrared sensor required a blind search (conducted by moving the end-effector through a predetermined pattern over a relatively large region) to acquire the target. Also, a single sensor provided only binary alignment data whereby the target is either sensed on-axis with the infrared beam or not sensed at all. Therefore, the implementation would have required the use of multiple sensors per target within the operating envelope to derive both translational and rotational error data. Second, the infrared triangulation methods offered less flexibility for target selection and identification. Third, the use of the triangulation method in the ASAL environment did not permit an operator to visually monitor operations and to provide assistance in conformity with the supervised autonomy approach. Consequently, emphasis was placed on the development of a machine vision sensor-based guidance system.

The basic elements of any machine vision system used for target identification are shown in sketch A. These elements form the foundation for the system developed for the automated assembly task. The video input, indicated at the top of the sketch, is the conventional scan signal from a video camera. The signal is input to an image processor that captures a single frame from the video signal and converts this information into a digital frame map. This frame map then is manipulated using a variety of standard image processing techniques that will be discussed. This information is passed to an algorithm that analyzes the processed image and identifies the desired target location in the two-dimensional image plane. These data are sent to a pose estimation algorithm to determine the location of the target (the target pose) in three-dimensional space relative to the vision system camera. The pose information, in the form of a linear homogeneous transformation matrix, then is output to the guidance control routine of the robot.



Sketch A

Machine Vision Requirements

Preliminary studies and tests using a machine vision system indicated that such a system could be designed to satisfy the constraints of small size, limited required servo region, incorporation of an operator-assisted mode, and to provide a logical means of discriminating a target from its surroundings. Based on these preliminary findings, the following set of requirements was used as a guide for the machine vision system development:

1. The camera and supporting equipment must
   - Be attached to the robot end-effector; the use of a stationary mounted system could not detect all positions required for strut installation
   - Be small and have low mass to permit non-obtrusive attachment to the robot end-effector
2. The system must
   - Enable an operator to monitor target identification operations and provide assistance consistent with the supervised autonomy approach
   - Be capable of providing position and range information from images taken along the approach path without moving the manipulator to locate the target
   - Operate reliably in an uncontrolled environment with minimal operator intervention
   - Be capable of discriminating between similar targets located on the same truss node
   - Be capable of rapidly acquiring and processing an image to avoid impeding the assembly process
   - Provide a positional accuracy of at least 0.25 in. to be consistent with the passive guidance features designed into the end-effector and truss assembly hardware
   - Not employ hazardous materials or hazardous energy sources to ensure operational safety
   - Require minimal maintenance and provide trouble-free operation
   - Offer self-calibration so that end-effector collisions do not disrupt system operation

Vision System Hardware Development

The vision system hardware includes a miniature video camera, the passive targets mounted on the joint receptacles, illumination system hardware, and a high-speed image processing system. The hardware required by the vision system includes a combination of off-the-shelf commercially available components selected for their unique features and specially designed and fabricated components that were not commercially available. Details of each of the hardware components are outlined in the following subsections.

Miniature Video Camera and Image Processing System

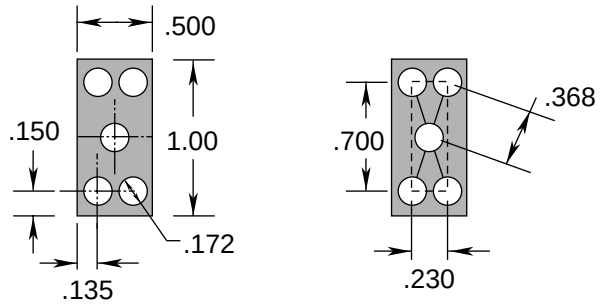
Several of the previously mentioned vision system requirements influenced the selection of the camera that was used. One major consideration was the limited space available for the location of the camera on the end-effector. The camera also had to be aligned with the target and be unobtrusive so that it did not affect the assembly scenario or end-effector component operations. Therefore, a miniature charge-coupled device (CCD) video camera similar to those mentioned in reference 1 was required to provide the video signal to the image processing system. Because most commercially available miniature video cameras do not have the capability for remote focusing and lens aperture setting, the range of operation and lighting conditions had to be bounded.

The image processing hardware consisted of a commercially available image processor controlled by a host computer. The video signal of the CCD

camera was sent directly to the image processing hardware where the image was captured, stored, and processed. Details on the various processes performed by this hardware are given in the section entitled "Image Processing Operations."

Target Definition and Fabrication

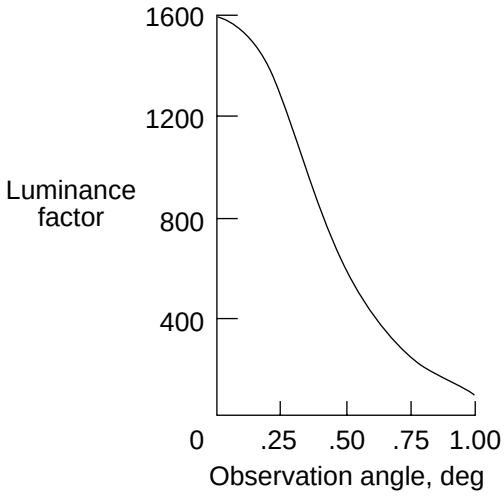
Joint receptacle targets were designed to satisfy a variety of requirements. Because the targets had to be mounted on the joint receptacles, they had to be small and not interfere with the capture of the joint receptacle by the end-effector fingers. The targets had to be simple, accurate, and capable of being mass-produced because hundreds were required. The use of active targets was rejected because they would have required synchronizing components and a power source for the target. Also, the targets had to provide information from which range and planar positioning could be determined so that the arm could be guided to the target in three-dimensional space. One target configuration that satisfied these requirements consisted of dots arranged in a distinctive pattern to facilitate target recognition by the image processing algorithm. Figure 2 shows a photograph of the joint receptacle target that measures approximately 0.5 in. by 1 in. Detailed dimensions (measured in inches) of the target design are shown in sketch B. A truss node that was partially populated with joint receptacles and which had some targets in place is shown in figure 3. For development tests, the targets were attached to the joint receptacles with a thin aluminum mounting platform. The proximity of the adjacent targets and the limited space available in the vicinity of the node illustrate the significance of the size constraint placed on the targets.



Sketch B

The target is constructed in layers to simplify fabrication. The bottom layer is a retroreflective tape, and the top layer is a thin sheet of flat-black anodized metal with five prepunched holes. The flat-black anodized metal mask provides a high contrast between the retroreflective target dots and the mask. This black mask and hole spacing also reduces the possibility of the dots blending with each other or with the image background. The domino configuration of five dots is simple enough to be accurately mass produced using a specially designed punching tool. This target pattern has been designed so that it can be easily discriminated from most background clutter, as described in the section entitled "Target Identification and Discrimination."

The retroreflective tape, which is an exposed lens retroflective plastic film precoated with a pressure-sensitive adhesive, is a High Gain Screen 7610 tape produced by 3M. This tape was selected because of its high reflective efficiency. Sketch C from 3M Products Bulletin "3M Special Effects Projection Screens" shows the reflective efficiency of the material in terms of the luminance factor as a function of the observation angle. The luminance factor is the number of times the tape is brighter than a perfect white diffuser. The observation angle is defined as the angle between the incident light ray (projector) and the reflected light ray at the receiver (camera). The target is most easily detected when it is actively illuminated by a light source that is as close to the optical axis of the camera as is possible to maximize the reflective efficiency of the target dots, as indicated by sketch C.



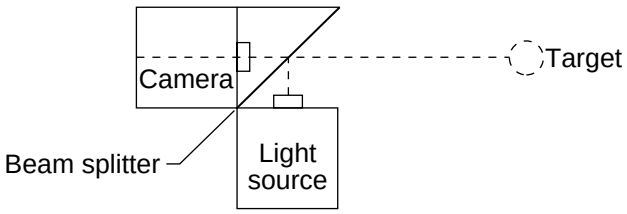
Sketch C

Light Assembly

The first-generation light assembly consisted of a simple aluminum ring centered around the CCD camera lens that held eight miniature incandescent light bulbs. Functionally, this setup provided a quick method to demonstrate the effectiveness of the machine vision system and the actively illuminated retroreflective target. However, such a simple setup had several drawbacks. First, the light source composed of the eight bulbs could not be placed on the optical axis of the camera, so an immediate loss of reflective efficiency was incurred. Second, the lights

produced a substantial amount of heat which could influence the image stability of the CCD array. Finally, the large amount of light source diffusion resulted in only a small amount of light reaching the target at ranges greater than 12 in., thus resulting in a low contrast when the target was backlit by external sources such as overhead lights and windows in the laboratory.

These drawbacks in conjunction with the vision system requirements outlined in the section entitled "Machine Vision Requirements" defined the basic form of the current camera and light assembly hardware, shown in figure 4. This current vision hardware, although larger than its predecessor, has several advantages that justify the additional size. The same CCD camera is used, but in place of the eight miniature incandescent bulbs are five commercially available miniature lights with focusing reflectors to concentrate the illumination (focusable flashlights). Four of the lights provide a focused off-axis light source, and the fifth, through the use of a partially silvered glass beam splitter, provides a light source on and aligned with the optical axis of the camera. Although a loss of intensity of the on-axis light source through the beam splitter occurs, this light still provides the greatest signal return from the target dots. The on-axis source is created by placing the light source at 90° to the camera lens and then setting the silvered glass at 45° between the camera and the light source, as shown in sketch D.



Sketch D

Integration With Robot End-Effector

The vision hardware mounted to the robot end-effector is shown in figure 5. Note that the beam splitter light is contained within the frame of the end-effector, and hence a reasonably low profile of 1.25 in. perpendicular to the axis of the end-effector is maintained. The current end-effector design does not permit the vision hardware to be centered and on axis with the end-effector fingers.

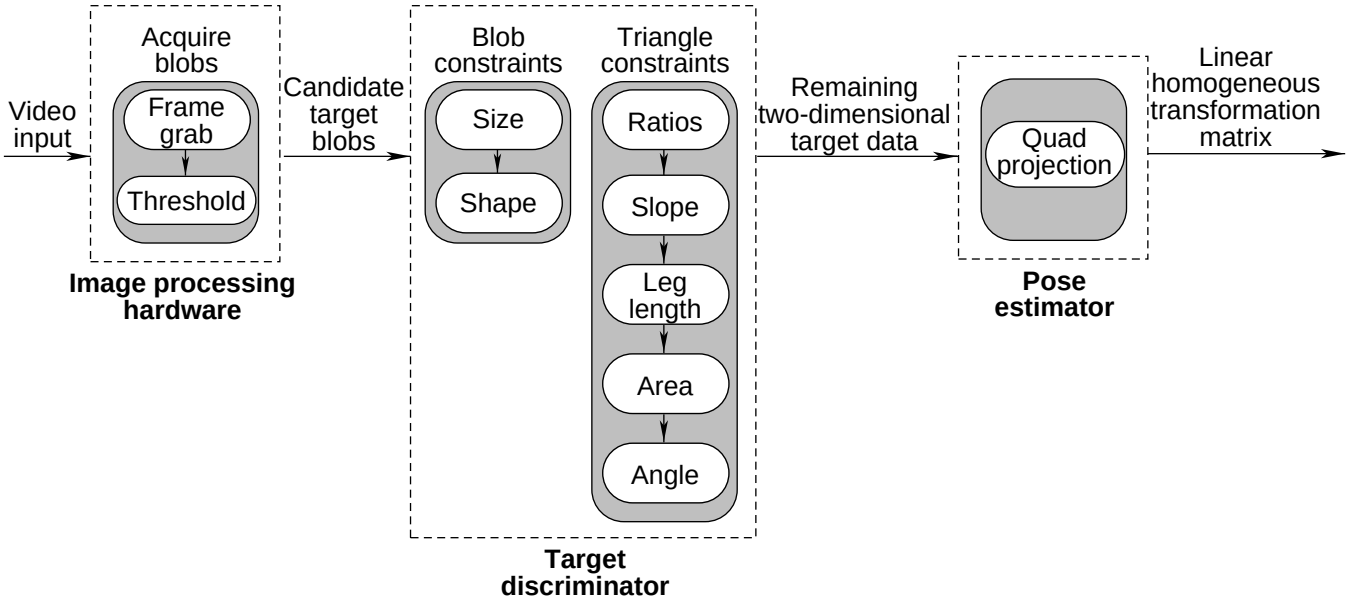
The robot arm at a typical VAP is shown in figure 6. From this perspective, the target mounted on the joint receptacle is not visible in the figure. The end-effector is approximately 18 in. away from the strut to be removed and pitched a few degrees to account for the offset of the vision hardware from the end-effector fingers.

The perspective from the video camera of a similar approach before illumination by the light assembly is shown in figure 7(a). The photograph on the left side of the figure is taken directly from the operator's monitor display, and it clearly shows that the target dots are not easily distinguished from their background, although the image is clear enough for the operator to locate the target if assistance is necessary. The data on the right side of the figure show the gray level (brightness of each picture element or pixel) array of the region highlighted in the figure. Note that the nominal value of pixels in the black background of the target periphery is approximately 77, and the nominal value of pixels in the target dots is approximately 106.

The same view illuminated by the light assembly is shown in figure 7(b). The direct effect of the actively illuminated reflective tape is evident. The gray level of approximately the same highlighted region is given in the data on the right of the figure, and it illustrates the improved contrast that an active lighting system provides. The nominal value of pixels in the background is approximately 81 (not a substantial increase over the previous figure), and the nominal value of pixels in the target dots is approximately 190, a substantial difference compared with the unilluminated target. The target dots in figure 7 appear noncircular in the data because of the aspect ratio of the CCD camera image array.

Image Processing Operations

The vision system hardware, discussed in the preceding sections, provides a standard video image containing the target shown in figure 2. The combination of retroreflective dots and active lighting allows discrimination of the target from its background through standard digital image processing techniques such as determining an image threshold, identifying feature centroids, and matching a model. The image processing elements consist of image processing hardware, candidate target "blob" discrimination software (the term blob refers to a contiguous region of similarly valued pixels), and pose estimation algorithms, as depicted in the image processing flowchart shown in sketch E. The video input to the processor hardware is a standard RS-170 video format. The output to the robot guidance control routine consists of a linear homogeneous transformation matrix that relates the camera lens center to the target. The image processing elements are discussed in detail in the following sections.



Sketch E

Video Image Processing

The image processing hardware, as mentioned in the section entitled "Miniature Camera and Image Processing System," consists of a commercially available image processor hosted by a computer. As configured, the processor can digitize up to eight RS-170 video signals, perform graphic generation and video display, provide gray level feature extraction and histograms at a frame rate of 30 frames per sec, and identify and analyze up to 255 blobs per frame every 33 msec.

The video signal is digitized into a 512×512 pixel frame. The gray level ranges from 0 (black) to 255 (white). The process of digitizing an image is commonly called frame grabbing. The digitized image is converted into a binary image by selecting a threshold value and assigning all the gray values that are less than that value to black (0) and all gray values that are equal or greater than that value to white (1). The threshold value is either user defined or calculated based on a histogram of the image. For the calculated threshold value, two gray-level percentage parameters are defined by the operator. One parameter specifies a percentage of expected background pixels (low intensity). The other parameter specifies a percentage of foreground (high intensity) pixels which should contain the target dots. The midpoint between the gray level values associated with these percentages is used as the threshold value. The image processing hardware then provides an array of candidate target blobs which may contain the target dots.

Target Identification and Discrimination

Discrimination of the target from the array of candidate target blobs is enhanced by target design, active illumination, and target pose constraints developed to be in concert with the assembly operation. A variety of software constraining techniques are applied to the array of candidate target blobs to identify the target dots within this array. The discrimination process block depicted in sketch E initially attempts to eliminate candidate target blobs based on their size and shape. The centroids of all blobs that pass this first test then are used as vertices of triangles that are compared with known conditions of the target model triangles. Vertex blobs forming candidate triangles are accepted or rejected based on a comparison of the following parameters with the target model triangles: leg length ratios, slope of the primary leg, projected leg lengths, area, and interior angles. The following sections detail the target identification process.

In the current laboratory test-bed, a large number of extraneous blobs resulting from overhead lighting and windows are removed with size and shape discriminators. Based on the targeting range, the size of candidate target blobs is constrained to be between 10 pixels and 650 pixels. In addition, the approximate range information provided by the robotic kinematics enables perspective projection to eliminate invalid blob sizes based on the pinhole camera model, as shown in equations (1a) and (1b). This process is optional and generally used only when other checks

fail to identify the target:

$$\text{Model } Pix_x = Tw_x \left(\frac{f}{d}\right)\left(\frac{512}{Sw_x}\right) \tag{1a}$$

$$\text{Model } Pix_y = Tw_y \left(\frac{f}{d}\right)\left(\frac{512}{Sw_y}\right) \tag{1b}$$

where f is the focal length of the camera, d is the distance between the target and the camera, Model Pix_x and Model Pix_y are the areas of the blob in pixels, Tw_x and Tw_y are the dimensions of the candidate target blob, and Sw_x and Sw_y are the sizes of the CCD camera sensors in the directions of the X- and Y-axes, respectively. The constant 512 is the size in pixels of the frame buffer array in the direction of both the X- and Y-axes. If the percent error between the pinhole model and the candidate target blob exceeds a specified tolerance ε_1, as indicated in equation (2), the blob is discarded:

$$\left|\frac{\text{Model } Pix - \text{Target } Pix}{\text{Model } Pix}\right| > \varepsilon_1 \tag{2}$$

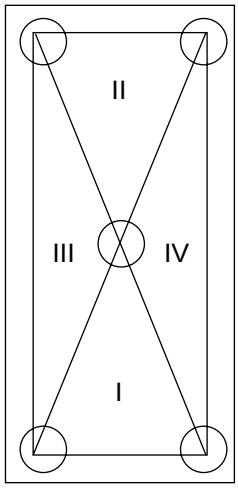
Because the five dots comprising the target are circular, distinctly noncircular blobs can be discarded. The following constraints ensure that highly elongated blobs are eliminated:

$$\max X - \left(\min X + 2\sqrt{\frac{n}{\pi}}\right) < \varepsilon_2 \tag{3a}$$

$$\max Y - \left(\min Y + 2\sqrt{\frac{n}{\pi}}\right) < \varepsilon_2 \tag{3b}$$

where $\min X$, $\max X$, $\min Y$, and $\max Y$ are the blob extrema pixel coordinates and n is the size of the blob in pixels. The parameter ε_2 is a user-specified error tolerance to accommodate slight rotations and effects of shadowing which may alter the centroid of a blob. Both constraints must be satisfied for a blob to continue to be considered as a candidate target.

After constraining the array of candidate blobs by these tests, a complete set of $t!/3!(t-3)!$ triangles is generated from the t remaining blobs. Knowledge of the target configuration permits the use of simple geometric constraints based on properties of similar triangles. The vision target is decomposed into four triangles, as shown in sketch F. In general, triangles I and II and triangles III and IV are not symmetric because of rotational nonalignment of the CCD camera and the target. Each triangle has the center blob as a common vertex.



Sketch F

Pertinent information from each triangle generated from the array of remaining blobs is stored in memory; the legs are ordered according to length (in pixels). The following three ratios r then are ordered and stored:

$$r1 = \frac{d3}{d1} \tag{4a}$$

$$r2 = \frac{d2}{d1} \tag{4b}$$

$$r3 = \frac{d3}{d2} \tag{4c}$$

where $d1, d2$, and $d3$ are the lengths of the ordered legs. The ratios of candidate triangles are compared with each of the four sets of triangle ratios that correspond to triangles I, II, III, and IV associated with the model. Candidate triangle ratios that fail to match any of the four model ratios within a specified tolerance are discarded. Otherwise, the candidate is classified as I, II, III, or IV as defined in sketch F.

For each triangle class, a primary leg is defined for slope determination; this slope is defined as inplane rotations. For classes I and II, the horizontal leg is the primary leg. For classes III and IV, the vertical leg is the primary leg. The slope of the primary leg is calculated and, if the difference between the calculated and expected values exceeds a user-specified error tolerance, the candidate triangle is discarded.

For classes I and II,

$$\left.\begin{array}{l} |m| = |\pi \pm \varepsilon_3| \\ \text{or} \\ |m| = |0 \pm \varepsilon_3| \end{array}\right\} \tag{5a}$$

and for classes III and IV,

$$\left|\frac{\pi}{2} - \varepsilon_3\right| \leq |m| \leq \left|\frac{\pi}{2} + \varepsilon_3\right| \qquad (5b)$$

where m is the slope of the primary leg and ε_3 is a specified error tolerance.

The approximate range provided by the robotic kinematics enables the use of perspective projection to eliminate invalid triangles based on the pinhole camera model described in equations (1a) and (1b). The length of the prime leg, in pixels, is compared with the pinhole projection at the supplied range. If the comparison exceeds a specified tolerance, the triangle is discarded.

The area of the acceptable triangles, as with blob size, is also bounded based on the known operating range. When the lengths of the legs of each triangle are known, Heron's formula, given in equation (6), can be used to determine the area of each triangle:

$$\text{Area} = \sqrt{k(k-a)(k-b)(k-c)} \qquad (6)$$

where a, b, c are the leg lengths, and $k = \frac{1}{2}(a+b+c)$. Triangles whose areas are outside a specified bounded region are discarded.

Each of the three interior angles in a triangle is also bounded based on the target model. For the current target configuration, the angles must be between 0.14 rad (8°) and 2.9 rad (166°). Triangles whose angles are outside this range are discarded. Those triangles that remain should have the target dots as their vertices. Prior to passing the remaining triangle vertices to the pose estimation algorithm, a verification check is imposed. This check verifies that the remaining triangles have a common vertex that lies somewhere between the other vertices. Although this check does not guarantee that the remaining triangle vertices are the target dots, the check does decrease the possibility that the vertices sent to the pose estimation algorithm are incorrect ones.

Pose Estimation

A quadrangle projection algorithm (refs. 3 to 7) is applied to determine the position and orientation of the target with respect to the camera lens center once the five target dots are identified. Because the algorithm requires a convex quadrilateral of known size, the centroids of the four outer points of the target are used as vertices.

Preliminary Test Results and Discussion

The machine vision system was not only tested on the ASAL robot but also laboratory tested on a vibration-free optical bench. In the laboratory tests, the camera and light assembly was rigidly mounted to the optical bench, and the target was attached to a six-degree-of-freedom articulator. The ASAL test was used to verify the target discrimination process in a noncontrolled environment. The optical bench test was used to verify the pose estimation algorithm in a controlled environment. The results of these tests follow.

ASAL Robot Tests

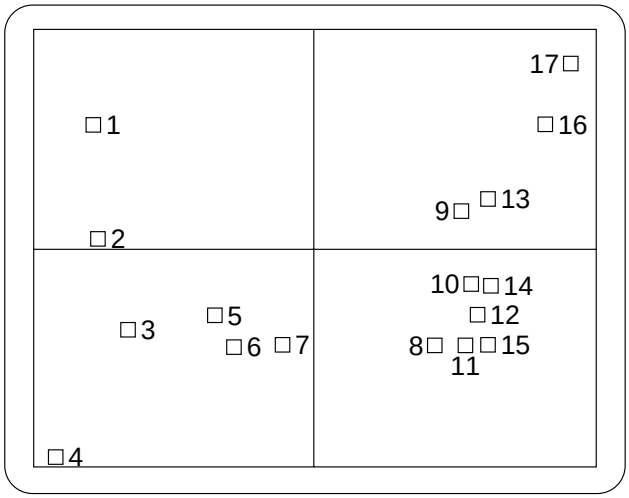
The vision system is designed to process images in the defined working envelope between 24 in. and 8 in. from the target. The 24-in. value takes into account the offset of the video camera from the end-effector fingers and any discrepancy between the approximate VAP of 18 in. and the actual VAP for a particular strut. Because the movement of the arm toward the target is done incrementally, the vision system also must be able to process intermediate-range images. To illustrate the types of images to be processed, one such incremental approach is shown in figures 8(a) to 8(d). The orientation of the video camera makes the strut appear to be upside down and rotated. The light at the bottom of the figure is sunlight shining through the open laboratory door. Figure 8(a) shows the image to be processed at a range of approximately 24 in. Each dot of the target is approximately five pixels in diameter, which approaches the limit of reliable processing. Figure 8(b) presents the image to be processed at a range of approximately 18 in. Each target dot is about seven pixels in diameter. Figure 8(c) shows the image to be processed at a range of approximately 12 in. Each target dot is approximately 10 pixels in diameter. Finally, figure 8(d) presents the image to be processed at a range of approximately 6 in. Each target dot is about 20 pixels in diameter. A detailed explanation of the processing techniques for one of these intermediate images is given in the following paragraphs.

A variety of typical approach positions were tested to verify the effectiveness of the target identification process. The following example demonstrates the performance of the machine vision system at the intermediate approach position (approximately 12 in.) shown in figure 8(c). The field of view of the camera contains considerable background clutter such as ceiling lights, reflections, and an open laboratory door.

This initial set of candidate target blobs, the centroids of which are shown highlighted by the white squares with dark centers in figure 9(a), already has

been subjected to a minimum and maximum blob-size constraint. For the working range of the vision system, the size of the candidate target blobs is constrained to be between 6 pixels and 650 pixels in size. The area within the large, white rectangle is defined to be the region of interest. Only blobs residing within this region are considered as candidate target blobs. The size and position of the region of interest can be adjusted by the operator if needed. Such adjustments usually are not necessary because the default size and position of the region are set to accommodate typical approach paths.

Sketch G enumerates the retained feature blobs shown in figure 9(a). These image blobs are referenced in the remainder of this section. To illustrate, blobs and triangles that fail their respective checks are erased, but blobs used as vertices of triangles that fail are not erased.



Sketch G

For this example, the use of equations (1) and (2) to eliminate blobs based on their projected size as a function of the distance between the camera and the target is not necessary. In general, ε_1 in equation (2) is between 0.1 and 0.5 depending on the distance between the camera and the target. To eliminate elongated or elliptical blobs, equations (3a) and (3b) are used. Figure 9(b) shows the centroids of the blobs that remain after they are screened by shape. For this example, blob 3 in sketch G is a typical blob that is generated by a reflection along the length of the strut with the following associated values: $\max X = 186$, $\min X = 156$, and $n = 46$. Given $\varepsilon_2 = 7$, the inequality from equation (3a) fails:

$$186 \not< \left(156 + 2\sqrt{\frac{46}{\pi}} + 7\right)$$

and

$$186 \not< 171$$

In contrast, blob 10, which is one of the five target dots, has the associated values $\min X = 345$, $\min Y = 220$, $\max X = 351$, $\max Y = 228$, and $n = 50$, thus resulting in the inequalities from equations (3a) and (3b):

$$351 < 360$$

and

$$228 < 235$$

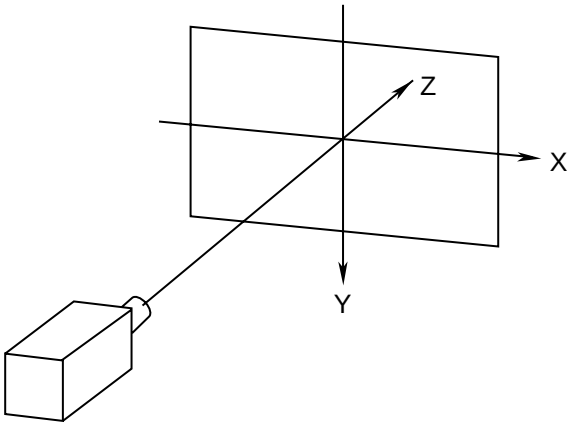
This process eliminates blobs 2, 3, 6, 7, and 8.

Figure 9(c) shows the set of triangles after exhaustive generation from the remaining blobs after size and shape screening. Given the list of the 12 remaining blobs in the figure for this particular approach position, 220 triangles are generated. Figure 9(d) shows the triangles remaining after the application of the ratio test equations (4a), (4b), and (4c). Note that many triangles have been eliminated but only blob 1 is eliminated by this test. The model ratios shown in sketch F for class I and II triangles are $r1 = 1.97$, $r2 = 1.97$, and $r3 = 1.00$, and the model ratios for class III and IV triangles are $r1 = 1.93$, $r2 = 1.93$, and $r3 = 1.00$. Given a specified tolerance of 0.10, the ratios for the triangle with vertices at blobs 1, 4, and 5 are $r1 = 1.7$, $r2 = 1.12$, and $r3 = 1.56$; this triangle fails the ratio test. The model ratios for the triangle with vertices at blobs 10, 11, and 12 are $r1 = 1.96$, $r2 = 1.91$, and $r3 = 1.02$; this triangle passes the test for classes III and IV.

Figure 9(e) shows the triangles remaining after application of the slope test. Note that two additional blobs and their associated triangles have been eliminated. For all vision approach points, the target will be rotated no more than 20° from the orientation shown in sketch F. Therefore, class I and II triangles must have a primary leg that is almost horizontal (0° or 180°), and class III and IV triangles must have a primary leg that is almost vertical (90°). Specifying $\varepsilon_3 = 0.35$ rad (20°) for equations (5a) and (5b), the primary leg of the triangle with vertices at blobs 4, 5, and 10, classified as both a III and IV triangle, has a slope of 0.478 rad (27.39°), and it is therefore discarded because it is greater than 20° from vertical. The triangle with vertices at blobs 10, 11, and 12, identified as a class III and IV triangle, has a slope of 1.52 rad (87.1°) and is retained. This test also is used to discriminate between multiple targets on a fully populated node because only one target consists of triangles that can pass the slope test because of the orientation of the camera with respect to the node.

Figure 9(f) shows the triangles remaining after application of the triangle area constraint. For this particular example, all remaining nontarget triangles are eliminated by the area check. For example, given a specified bounded triangle region of 10 pixels to 750 pixels, the triangle with vertices at blobs 13, 14, and 16 fails with an area of 827 pixels, but the triangle with vertices at blobs 10, 11, and 12 passes with an area of 98 pixels.

The cross hair located on the center target dot in figure 9(g) indicates that the target was successfully located. Once located, the pose estimation algorithm is invoked, and the results are both displayed on the monitor (fig. 9(g)) and relayed to the robot guidance and control routine. The numbers printed in figure 9(g) show the translation and orientation of the example approach point. The top two lines give the position vector in both inches and millimeters, respectively. The bottom line gives the rotation vector in radians. Sketch H denotes the camera axis system for these vectors.



Sketch H

As a result of testing the machine vision system in ASAL, values for each of the software target identification parameters were derived empirically. The following set of parameter values in the table enabled target acquisition for all approach positions attempted in preliminary tests:

Minimum blob size, pixels	6
Maximum blob size, pixels	650
Minimum triangle area, pixels	10
Maximum triangle area, pixels	750
Minimum triangle internal angle, radians	0.14
Maximum triangle internal angle, radians	2.9
Circular blob tolerance, pixels	7
Triangle leg ratio tolerance	0.10
Triangle rotation tolerance, radians	0.35
Utilization of camera distance from target	Disabled
Automated threshold parameters, percent background pixels	91
Automated threshold parameters, percent foreground pixels	1.5

Optical Bench Tests

The graphs in figures 10(a) to 10(c) show the translational accuracy of the pose estimation for each of the three axes. The optical bench test results for an absolute position error along the vision system X-axis (sketch H) are shown in figure 10(a). The plot shows the absolute position error along the X-axis as a function of the relative horizontal position of the target within the region of interest where the region is subdivided into 10 equal divisions. Results are presented for the Z-axis camera positions of 8 in., 12 in., 16 in., and 20 in. The physical width of each region of interest at the specific range value is given in parentheses beside the corresponding range value (i.e., at a range of 8 in., the region of interest is 3.5 in. wide, and each horizontal division is 0.35 in.). The zero relative position within the region of interest defines the horizontal location of the optical axis of the camera. Data collection for each range value begins at the optical axis; the articulator is adjusted until the range determined by the pose estimation routine matches the measured range. Subsequent values then are taken after translating the articulator in the X-axis direction. The plot shows that all but one data point fell significantly below the maximum allowable position error of 0.25 in. The majority of the data fell within 0.05 in. of the actual X-axis position. The tendency of the absolute position error to increase toward the extremes of the region of interest is probably attributed to radial distortion inherent in the small focal length (7.5 mm) camera lens. The variation in the absolute position error across the region of interest may be partially because of variations in light intensity of the active illumination source. Such a phenomenon causes variation in the gray level intensity of the target dots which can lead to errors in determining the centroid of the target dots.

The optical bench test results for absolute position error along the vision system Y-axis are shown in figure 10(b). The plot shows the absolute position error along the Y-axis as a function of the relative vertical position of the target within the region of interest where the region is subdivided into 10 equal divisions. Results are presented for the Z-axis camera positions of 8 in., 12 in., 16 in., and 20 in. The physical height of each region of interest

at the specific range value is given in parentheses beside the corresponding range value (i.e., at a range of 8 in., the region of interest is 3.0 in. high, and each division is 0.3 in.). The zero relative position within the region of interest defines the vertical location of the optical axis of the camera. Data collection for each range value begins at the optical axis; the articulator is adjusted until the range determined by the pose estimation routine matched the measured range. Subsequent values then are taken by translating the articulator in the Y-axis direction. The plot shows that all but one data point fell below the maximum allowable position error of 0.25 in. The majority of the data fell within 0.05 in. of the actual Y-axis position with slightly more points above 0.05 in. than in the previous plot. This slight increase in overall position error between the points along the Y-axis versus those along the X-axis and the more pronounced tendency of the absolute position error to increase toward the extremes of the region of interest are probably indicative of an increased effect of lens distortion along the vertical axis. The data presented are raw data, and no corrections have been made for lens distortion.

The optical bench test results for the vision system determined range are shown in figure 10(c). The plot shows the vision system range data as a function of the measured range. The target was centered on the optical axis of the camera to minimize lens distortion effects. Between the 8-in. and 18-in. ranges, the results indicate good correlation with measured range values. Beyond 18 in., the discrepancy between the reported value and the measured value becomes more prominent. Pixel resolution may be largely responsible for the increased error measured at the larger range values. The results of figures 10(a) to 10(c) indicate that the vision system is capable of providing the accuracies required by ASAL.

Future Research Opportunities

This machine vision system has demonstrated the potential to significantly enhance the capability of the ASAL operations. The vision system and a path planner provide the opportunity to eliminate the reliance on taught points for assembly and disassembly operations, thereby simplifying the current and any new truss assembly procedures. The accuracy of the system may be enhanced by performing a rigorous calibration of both the CCD camera and the system as a whole.

Summary of Results

An ongoing research program at the NASA Langley Research Center Automated Structural Assembly Laboratory (ASAL) is evaluating a telerobotic automated structural assembly system designed to assemble large space truss structures. This system relies on predetermined ("taught") points for truss structural assembly. Reliance on these taught points for joint receptacle capture and strut installation is neither robust nor reliable enough for space operations. Therefore, work on a machine vision sensor guidance system has been initiated to develop and evaluate the capability for locating and guiding the robot to a passive target mounted on the joint receptacle.

The hardware portion of this sensor guidance system consists of a miniature charged-coupled device (CCD) camera and light assembly mounted on the robot end-effector and a target mounted on each joint receptacle. The camera sends its video signal to a remotely located image processor. The target, which measures approximately 0.5 in. by 1 in., consists of a flat-black background with five retroreflective dots. The dots, which are arranged in a distinctive pattern to facilitate target recognition by an image processing algorithm, are actively illuminated by five miniature lights with focusing reflectors. Four of the lights provide a concentrated off-axis light source and the fifth, through the use of a partially silvered glass beam splitter, provides a light source directly on the optical axis of the camera. By maintaining a light source on the optical axis of the camera, the reflective efficiency of the target dots is maximized.

Discrimination of a target from background clutter is accomplished through standard digital image processing techniques such as determining an image threshold, identifying feature centroids ("blobs"), and matching a model. (The word blob is defined as a contiguous region of similarly valued pixels.) Determining an image threshold, based on a histogram of the image, followed by identifying the region centroid, provides a list of candidate target blobs. An image processing algorithm has been developed to discriminate the target dots from this list of candidate target blobs. The algorithm first screens blobs according to their size and shape, retaining only those blobs that are within a defined tolerance band of the target model requirements. The centroids of the remaining blobs are used to form a complete set of triangles which is subjected to a variety of tests based on geometric features of the target model triangles. Eliminating triangles that do not satisfy constraints of the target model triangles further reduces the number of candidate target blobs. Once successfully discriminated, the centroids of the remaining five target dots are sent to a pose estimation routine to determine the location, in three-dimensional space, of the target relative to the robot end-effector.

Preliminary test results of the vision system in the ASAL facility with a range of lighting and background conditions indicate that the system is flexible enough to successfully identify joint receptacle targets throughout the required operational range of the system. Controlled optical bench tests of the vision system indicate that the system is fully capable of providing the accuracies required for pose estimation of the target.

NASA Langley Research Center
Hampton, VA 23665-5225
March 31, 1992

References

1. Rhodes, Marvin D.; Will, Ralph W.; and Wise, Marion A.: *A Telerobotic System for Automated Assembly of Large Space Structures.* NASA TM-101518, 1989.

2. Wu, K. Chauncey; Adams, Richard R.; and Rhodes, Marvin D.: *Analytical and Photogrammetric Characterization of a Planar Tetrahedral Truss.* NASA TM-4231, 1990.

3. Haralick, Robert M.: Using Perspective Transformations in Scene Analysis. *Comput. Graph. & Image Process.,* vol. 13, no. 3, July 1980, pp. 191–221.

4. Myers, Donald R.; Juberts, Maris; and Leake, Stephen A.: Enhanced Telemanipulator Operation Using a Passive Vision System. *IEEE 1985 Proceedings of the International Conference on Cybernetics and Society,* IEEE Catalog No. 85CH2253-3, IEEE Systems, Man, and Cybernetics Soc., 1985, pp. 802–806.

5. Goode, Plesent W.; and Cornils, Karin: Monovision Techniques for Telerobots. Paper presented at the Workshop on Space Telerobotics (Pasadena, California), Jan. 20–22, 1987.

6. Cornils, Karin; and Goode, Plesent W.: *Location of Planar Targets in Three Space From Monocular Images.* NASA TM-101868, 1987.

7. Hung, Yubin; Yeh, Pen-Shu; and Harwood, David: Passive Ranging to Known Planar Point Sets. *1985 IEEE International Conference on Robotics and Automation,* IEEE Catalog No. 85CH2152-7, IEEE Computer Soc., 1984, pp. 80–85.

(a) Schematic of Automated Structural Assembly Laboratory.

(b) Hardware in Automated Structural Assembly Laboratory.

Figure 1. Automated Structural Assembly Laboratory.

L-90-5053

L-92-19

L-91-6330

Figure 2. Joint receptacle target.

Figure 3. Truss node with joint receptacle targets.

L-91-6331

Figure 4. Camera and light assembly.

L-91-05738

Figure 5. End-effector mounted camera and light assembly.

L-92-20

Figure 6. Robot arm at typical vision approach point.

L-91-6379

(a) Before illumination.

Figure 7. CCD camera image with and without active lighting and corresponding target region gray scale.

(b) After illumination.

Figure 7. Concluded.

L-91-6380

(a) Vision system view at range of $\approx$24 in.

(b) Vision system view at range of $\approx$18 in.

Figure 8. Typical incremental approach.

L-91-6367

L-91-6366

L-91-6368

L-91-6369

(c) Vision system at range of $\approx$12 in.

(d) Vision system view at range of $\approx$6 in.

Figure 8. Concluded.

L-91-6370

L-91-6371

(a) Processed image after blob-size constraint applied.

(b) Processed image after blob shape constraint applied.

Figure 9. Sequential discrimination of target from background.

L-91-6372

L-91-6373

(c) Processed image after exhaustive triangle generation.

(d) Processed image after application of ratio test equations.

Figure 9. Continued.

L-91-6374

L-91-6377

(e) Processed image after triangle slope constraint applied.

(f) Processed image after length, area, and angle constraints applied.

Figure 9. Continued.

L-91-6378

(g) Processed image after target lock-on and pose estimation. Top two lines of key give position vector in inches and millimeters, respectively. The bottom line gives rotation vector in radians.

Figure 9. Concluded.

(a) Absolute position error along vision system X-axis.

Figure 10. Vision system results from optical bench tests for Z-axis camera position. Positions are measured in inches. (Numbers in parentheses in key indicate region width in inches.)

(b) Absolute position error along vision system Y-axis.

Figure 10. Continued.

(c) Vision system range data versus measured range.

Figure 10. Concluded.